Centre for Research and Analysis of Migration

CReAM

Discussion Paper Series

CDP No 07/07

International migration: A panel data analysis of the determinants of bilateral flows

Anna Maria Mayda

Centre for Research and Analysis of Migration
Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

CReAM Discussion Paper No 07/07

International migration: A panel data analysis of the determinants of bilateral flows

Anna Maria Mayda†

† Department of Economics and School of Foreign Service,
Georgetown University.

Non-Technical Abstract

In this paper I empirically investigate the determinants of migration inflows into fourteen OECD countries by country of origin, between 1980 and 1995. I analyze the effect on migration of average income and income dispersion in destination and origin countries. I also examine the impact of geographical, cultural, and demographic factors as well as the role played by changes in destination countries' migration policies. My analysis both delivers estimates consistent with the predictions of the international migration model and generates empirical puzzles.

JEL classification: F22, F1.
Keywords: International Migration, Determinants, Push and Pull Factors, Migration Policy.

Centre for Research and Analysis of Migration
Department of Economics, Drayton House, 30 Gordon Street, London WC1H 0AX
Telephone Number: +44 (0)20 7679 5888
Facsimile Number: +44 (0)20 7916 2775

International migration:

A panel data analysis of the determinants of bilateral flows*

Anna Maria Mayda†

April 2007

Abstract

In this paper I empirically investigate the determinants of migration inflows into fourteen OECD countries by country of origin, between 1980 and 1995. I analyze the effect on migration of average income and income dispersion in destination and origin countries. I also examine the impact of geographical, cultural, and demographic factors as well as the role played by changes in destination countries' migration policies. My analysis both delivers estimates consistent with the predictions of the international migration model and generates empirical puzzles.

JEL classification: F22, F1.

Keywords: International Migration, Determinants, Push and Pull Factors, Migration Policy.

*This paper was first circulated with the title "International migration: A panel data analysis of economic and non-economic determinants," May 2003. I would like to thank Alberto Alesina, Elhanan Helpman, Dani Rodrik, Klaus Zimmermann for many insightful comments. For helpful suggestions, I am also grateful to Richard Adams, Marcos Chamon, Giovanni Facchini, Bryan Graham, Louise Grogan, Russell Hillberry, Arik Levinson, Lindsay Lowell, Rod Ludema, Lant Pritchett, Maurice Schiff, Tara Watson, Jeffrey Williamson, and participants in the International Workshop at Harvard University, at the 2003 NEUDC Conference at Yale University, at the International Trade Commission Workshop, at the World Bank International Trade Seminar and at IZA Annual Migration Meeting. I am grateful for research assistance provided by Krishna Patel and Pramod Khadka. All errors remain mine.

†Department of Economics and School of Foreign Service, Georgetown University, ICC 552, 37th and O Streets, NW, Washington DC, 20057; tel 202 687 6712; fax 202 687 6102; email: amm223@georgetown.edu.

1 Introduction

International migration patterns vary considerably over time, and across destination and origin countries. Some OECD countries have experienced a decrease in the size of the annual immigrant inflow between 1980 and 1995.[1] Over the same years, the number of immigrants per year has increased in several other OECD countries.[2] The percentage change of the annual immigrant inflow from 1980 to 1995 ranges between negative 42% (in Japan) and positive 48% (in Canada) (OECD 1997). For all destinations, such changes are anything but monotonic (see Figure 1). The variation in terms of origin countries is remarkable as well (see Appendix 1).

Several factors are likely to influence the size, origin, and destination of labor movements at each point in time and contribute to the variation observed in the data. However, very few empirical works in the literature have tried to understand what drives international migration, perhaps due to past unavailability of cross-country data.

In turn, international migration has recently received a great deal of attention in light of research showing its beneficial effects from an economic-development point of view. For example, the recent literature has pointed out repeatedly the potential of free migration to produce large benefits – most likely greater than the gains from liberalizing existing trade barriers.[3] Other studies have uncovered the role played by foreign remittances of international migrants in their origin countries' economies.[4] To fully understand these and other effects, it is important to identify the forces and constraints that shape international migration movements.

In this paper, I empirically investigate the determinants – economic, geographic, cultural and demographic – of bilateral immigration flows. My analysis is based on the predictions of a simple theoretical framework that focuses on both supply and demand factors. I use yearly data on immigrant inflows into fourteen OECD countries by country of origin, between 1980 and 1995. The source of this data is the International Migration Statistics for OECD countries (OECD 1997), based on the OECD's Continuous Reporting System on Migration (SOPEMI).

My paper is related to a vast literature on the determinants of migration which includes works dating back to the nineteenth century (Ravenstein 1885). More recently, Clark, Hatton and Williamson (2002) and Karemera, Oguledo and Davis (2000) both focus on the fundamentals explaining immigrant inflows into the United States by country of origin in the last decades. Other papers in the literature that analyze the determinants of migration to the U.S. are Borjas (1987), Borjas and Bratsberg (1996) and Yang (1995). Hatton (2005) investigates trends in UK net migration in the last decades. Finally, Helliwell (1997, 1998) sheds light on factors affecting labor movements in his investigation of the magnitude of immigration border effects, using data on Canadian interprovincial, US interstate and US-Canada cross-border immigration.

This paper makes three contributions to the literature. First, my analysis puts greater

[1] For example, France, Japan, Netherlands, and the United Kingdom (OECD 1997).

[2] For example, in Belgium, Canada, Germany, Luxembourg, Norway, Switzerland, and the United States (OECD 1997).

[3] See Rodrik 2002, Pritchett 2003, Martin 2004.

[4] See, for example, Hanson and Woodruff 2004.

emphasis than previous works on the demand side of international migration, namely destination countries' migration policies. This change of perspective is important, given restrictive immigration policies in the vast majority of host countries. Second, my work is the first one I am aware of to use the OECD (1997) data on international migration to systematically investigate the drivers of international flows of migrants. Previous works have either used country cross-sections (Borjas 1987, Yang 1995), or have focused on a single destination country over time (Borjas and Bratsberg 1996, Clark, Hatton and Williamson 2002, Karemera, Oguledo and Davis 2000, Brücker, Siliverstovs, and Trübswetter 2003) or a single origin country over time (Yang 2003).[5] By extending the focus of the analysis to a multitude of origin and destination countries and taking advantage of both the time-series and cross-country variation in the data, I can test the robustness and broader validity of the results found in earlier works.[6] Third, this paper carefully reviews and proposes solutions to various econometric issues that arise in the estimation, such as endogeneity and reverse causality. These econometric complications have not all been addressed in the previous literature.[7] Once I deal with them (for example, by controlling for destination and origin countries' fixed effects), my analysis both delivers estimates broadly consistent with the predictions of the international migration model and generates empirical puzzles.

According to the international migration model, pull and push factors have either similar-sized effects (with opposite signs), when migration quotas are not binding, or they both have no (or a small) effect on emigration rates, when migration quotas are binding. It is not clear, ex ante, which one of the two scenarios characterizes actual flows. Migration policies in the majority of destination countries are very restrictive, which should imply binding constraints on the number of migrants. On the other hand, even countries with binding official immigration quotas often accept unwanted (legal) immigration.[8] Restrictive immigration policies are often characterized by loopholes, that leave room for potential migrants to take advantage of economic incentives. For example, immigration to Western European countries still took place after the late Seventies, despite the official closed-door policy (Joppke 1998). Family-reunification and asylum-seekers policies can explain continuing migration inflows to Western Europe.[9]

My empirical results are puzzling because they are in part consistent with the first scenario and in part with the second one. I find that pull factors - that is, improvements in the mean income opportunities in the destination country - significantly increase the size of emigration rates. This result is very robust to changes in the specification of the empirical model. Both absolute and relative pull factors matter. That is, the emigration rate to a given destination is an increasing function of that country's per worker GDP and a decreas-

[5] The paper most related to mine is Clark, Hatton, and Williamson (2002).

[6] Since I began working on this paper, I have become aware of other related, but independent papers analyzing cross-country migration patterns: Alvarez-Plata, Brücker, and Siliverstovs (2003), Pedersen, Pytlikova, and Smith (2004), Pedersen, Pytlikova, and Smith (2006). I discuss these very recent contributions to the literature below, in relation to the data I use and results I find in this paper.

[7] See also Alvarez-Plata, Brücker, and Siliverstovs (2003) for an excellent discussion of the properties of different estimators of the determinants of migration flows.

[8] Notice that the data set I use only covers legal migration.

[9] Joppke (1998) writes about Germany's experience (p.285): "Since the recruitment stop of 1973, the chain migration of families of guest workers was (next to asylum) one of the two major avenues of continuing migration flows to Germany, in patent contradiction to the official no-immigration policy."

ing function of the average per worker GDP of all the other host countries in the sample[10] (each weighted by the inverse of distance from the origin country). On the other hand, the sign of the impact of push factors - that is, declining levels of per worker GDP in the origin country - is seldom negative as theory suggests would be the case with not-binding migration quotas and, when it is, the size of the effect is smaller than for pull factors and is almost always insignificant. Therefore my analysis finds evidence of an asymmetric impact of pull and push factors on emigration rates.[11]

An interpretation that reconciles the results on positive and significant pull effects and small or insignificant push effects is that migration quotas are effectively not binding but the impact of income opportunities in the origin country is affected by poverty constraints, due to fixed costs of migration and credit-market imperfections (Lopez and Schiff 1998, Yang 2003). Since lower levels of per worker GDP in the source country both strengthen incentives to leave and make it more difficult to overcome poverty constraints, the net effect might be close to zero. In the empirical analysis I investigate this possibility and I find only weak evidence that my result on push factors is driven by poverty constraints in the origin country.

Yet an alternative explanation of my findings is that the asymmetric effect I estimate for pull and push factors is explained by the demand side of international migration - namely, migration policies - and not by the supply side as is often assumed in the previous literature. Changes in mean income opportunities in the destination country not only affect migrants' incentive to move there but also impact the political process behind the formation of migration policies. For example, in periods of economic booms, policymakers are better able to overcome political opposition to and accomodate increasing migration inflows.[12] If migration quotas are binding, the latter political-economy channel will be at work while the determinants on the supply side will have no (or a small) impact. This would explain the asymmetric effect I estimate for pull and push factors. While I do not investigate this interpretation directly[13], I find evidence which is consistent with migration policy playing a constraining role. In the empirical analysis, I differentiate the effect of pull and push factors according to changes in destination countries' migration policy. I find that the effect of pull

[10] Since the host countries in the sample receive a large fraction of immigrants in the world, it is not overly restrictive to focus on them. For example, according to the United Nations (2004), the list of leading host countries of international migrants in 2000 – as measured by the percentage of the world's migrant stock in each of these countries – includes the United States (20%), Germany (4.2%), France (3.6%), Canada (3.3%), Australia (2.7%), United Kingdom (2.3%), Switzerland (1%), Japan (0.9%), and the Netherlands (0.9%) (see Table ii.3, p.30). These countries all belong to my sample.

[11] This result is consistent with the findings in the literature on internal migration. See, for example, Hunt (2006), which provides an explanation of the result by breaking down data by age group: Origin region's unemployment rates (push factor) have an insignificant impact on migration flows because the insignificant effect for the young – who are not as sensitive to their own layoffs as the old – dominates the significant positive effect for the old. See also the related literature on internal migration referenced in Hunt (2006) (footnote 4).

[12] Hanson and Spilimbergo (2001) focuses on US border enforcement and shows that enforcement softens when the sectors that use illegal immigrants expand, which is evidence that migration policy is affected by changes in economic conditions in the destination country.

[13] This interpretation goes beyond the theoretical model in this paper, which assumes exogenous migration quotas. The empirical analysis of the endogenous determination of migration policy and its role in explaining the asymmetric effect of pull and push factors is outside the scope of this paper.

factors becomes more positive and the impact of push factors turns negative in those years when a host country's immigration laws become less restrictive. This is also true for the impact of other supply-side determinants such as geography and demographics (see below). In sum, my results suggest that migration quotas matter as they mitigate supply-side effects.[14]

My empirical analysis also finds that inequality in the source and host economies is related to the size of emigration rates as predicted by Borjas (1987) selection model. An increase in the origin country's relative inequality has a non-monotonic effect on the size of the emigration rate: the impact is estimated to be positive if there is positive selection, negative if there is negative selection. Among the variables affecting the costs of migration, distance between destination and origin countries appears to be the most important one: Its effect is negative, significant and steady across specifications. On the other hand, there is no evidence that cultural variables related to each country pair play a significant role. Demographics - in particular, the share of the origin country's population who is young - shape bilateral flows as predicted by the theory. Since the effect of geography and demographics works through the supply side of the model, their impact should be even stronger when migration quotas are relaxed, which is what I find in the data.

Finally, I empirically investigate the importance of network effects. Since immigrants are likely to receive support from other immigrants from the same origin country already established in the host country, they will have an incentive to choose destinations with larger communities of fellow citizens. Network effects imply that bilateral migration flows are highly correlated over time, which is what the data shows. However, it is not clear how to interpret this result. While it is consistent with supply factors (that is, network effects), it could also be driven by demand factors (family reunification policies, for example).

The rest of the paper is organized as follows. Section 2 presents a simple model of international migration. In Section 3 I describe the data sets used, while in Section 4 I discuss the estimating equations and some econometric issues that complicate the analysis. Finally, I present the main empirical results and robustness checks in, respectively, Section 5 and Section 6. Section 7 concludes.

2 Theoretical framework

Both supply and demand factors affect international migration flows. Migrants' decisions to move, according to economic and non-economic incentives, shape the supply side of labour movements. The host country's immigration policy represents the demand side, namely the demand for immigrants in the destination country. The theoretical framework in this paper is closely related to the previous literature (Borjas 1999a, Clark, Hatton and Williamson 2002), the main difference being the greater emphasis in my model on destination countries' immigration policy. I consider two countries: country 0, which is the origin of immigrant flows and country 1, which is the destination. I first focus on the supply side of immigration and

[14] This result is consistent with the findings in Hatton (2004) where emigration from Britain in the era of free migration (before 1914) is compared to emigration in 1950 onwards, when immigration policies were in place in the four main host countries of British migrants. The paper finds that, from the mid 1960s, the impact of economic and demographic forces "became less powerful as they were increasingly inhibited by immigration policies in the principal destination countries." (p.1)

look at the probability that an individual chosen randomly from the population of country 0 will migrate to country 1. In each country, wages are a function of the individual skill level (s_i). The wages that individual i receives in country 0 and would receive if he migrated to country 1 are respectively equal to $w_{0i} = \alpha_0 + \theta_0 \cdot s_i + \epsilon_{0i}$ and $w_{1i} = \alpha_1 + \theta_1 \cdot s_i + \epsilon_{1i}$, where the two disturbances have zero means over the origin country's population. In light of the empirical analysis below, based on aggregate data, it is helpful to rewrite individual i's wages in the two locations as a function of first and second moments of the income distributions (of the origin country's population) at home and abroad respectively:

$$w_{0i} = \mu_0 + v_{0i}, \text{ where } v_{0i} \sim N(0, \sigma_0^2), \tag{1}$$
$$w_{1i} = \mu_1^0 + v_{1i}, \text{ where } v_{1i} \sim N(0, \sigma_1^2), \tag{2}$$

where the correlation coefficient between v_{0i} and v_{1i} equals ρ_{01}, μ_0 equals $\alpha_0 + \theta_0 \cdot \overline{s}_0$ and μ_1^0 equals $\alpha_1 + \theta_1 \cdot \overline{s}_0$ ($\overline{s}_0$ is the mean skill level of the origin country's population). Notice that μ_1^0, which is equal to the mean wage of the origin country's population if it all migrated to country 1, is different from $\mu_1 = \alpha_1 + \theta_1 \cdot \overline{s}_1$, which is equal to the mean wage of the destination country's population in country 1 ($\overline{s}_1$ represents the mean skill level of the destination country's population). This point will be relevant in one of the robustness checks in the empirical analysis.

I assume that each individual has Cobb Douglas preferences for the two goods produced in the world (x_A and x_B),[15] which implies an indirect utility (function) from having an income y given by $v(p_A, p_B; y) = \overline{A}(p_A, p_B) \cdot y$.[16] I assume that each country is a small open economy characterized by free trade with the rest of the world:[17] therefore goods' prices p_A and p_B, as well as $\overline{A}(p_A, p_B)$, are given and equal across countries.[18] An individual in country 0 will migrate to country 1 if the utility of moving is greater than the utility of staying at home that is, given the assumptions above, if the expected income in country 1 net of migration costs is greater than the expected income in country 0. Following the literature, I can define an index I_i that measures the net benefit of moving relative to staying at home for a risk-neutral individual i:[19]

$$I_i = \eta_{01} \cdot w_{1i} - C_i - w_{0i}, \tag{3}$$

where η_{01} is the probability that the migrant from country 0 will be allowed to stay in country 1, and $C_i = \mu_C + v_i^C$, with $v_i^C \sim N(0, \sigma_C^2)$, represents the level of individual migration costs.[20] The correlation coefficients between v_i^C and (v_{0i}, v_{1i}) are equal to (ρ_{0C}, ρ_{1C}). The implicit assumption in (3) is that, if the migrant moves to but is not allowed to stay in the

[15] Preferences are therefore of the following form: $U(x_A, x_B) = A x_A^{1-\delta} x_B^{\delta}, \ \ 0 < \delta < 1, \ \ A > 0$.

[16] In this expression: $\overline{A}(p_A, p_B) = A(\frac{1-\delta}{p_A})^{1-\delta}(\frac{\delta}{p_B})^{\delta}$.

[17] Given free trade, what explains the difference in rates of return to labor across countries? The answer is that, besides free trade, the other conditions for factor-price-equalization are not satisfied: for example, if international productivity differences exist (Trefler 1993), then only *adjusted* factor-price-equalization holds.

[18] In the empirical analysis I adjust for international differences in goods' prices, using PPP income levels.

[19] The index I_i does not include a capital-income term because capital is assumed to be internationally mobile (and therefore rates of return to capital are equalized across countries).

[20] I assume that each individual knows the wage levels w_{1i} and w_{0i} he would get in each location, the migration costs C_i and the probability η_{01}.

destination country, he still incurs the migration costs C_i and gives up the home wage w_{0i}. In other words, the individual migrates to the host country before knowing whether he will be able to stay (for a longer period of time) and gain the income w_{1i}.[21] Immigrants may not be able to stay in the host country because of quotas due to a restrictive immigration policy.

The probability that an individual chosen randomly from the population of the origin country will migrate from country 0 to country 1 therefore equals:

$$P = \Pr[I_i > 0] = \Pr[\eta_{01} \cdot (\mu_1^0 + v_{1i}) - (\mu_C + v_i^C) - (\mu_0 + v_{0i}) > 0], \tag{4}$$

which can be rewritten as $P = 1 - \Phi(z)$, where $z = -\frac{(\eta_{01} \cdot \mu_1^0 - \mu_0 - \mu_C)}{\sigma_v}$, σ_v is the standard deviation of $(\eta_{01} \cdot v_{1i} - v_{0i} - v_i^C)$, and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal.[22] The probability in (4) is the *supply* emigration rate $\frac{I_{01}^S}{P_0}$, where I_{01}^S represents the size of the migration flow as determined by the supply side of the model and P_0 the population in the origin country.

Next, I assume that the destination country's immigration policy sets quantity constraints for immigrants coming from each origin country. Let I_{01}^D be the maximum number of migrants from country 0 allowed each year into country 1. These immigration quotas, which represent country 1's demand for immigrants from country 0, may or may not be binding. Only in the latter case does the emigration rate we observe in the data $(\frac{I_{01}}{P_0})$[23] equal the *supply* emigration rate $\frac{I_{01}^S}{P_0}$ defined above. On the other hand, if quantity constraints are binding, $\frac{I_{01}}{P_0}$ will be less than $\frac{I_{01}^S}{P_0}$. In general, the emigration rate we observe in the data is equal to the minimum of $\frac{I_{01}^S}{P_0}$ and $\frac{I_{01}^D}{P_0}$, and is represented in Figure 2 by the heavy lines, as a function of μ_1^0 and μ_h, $h = 0, C$. The figure assumes that quotas I_{01}^D are exogenous, which means that they are not affected by μ_1^0 nor by μ_h, $h = 0, C$. This is a strong assumption that is questioned in the interpretation of the empirical results.

I assume that the probability η_{01} that the migrant from country 0 will be allowed to stay in country 1 is equal to $\min\{1, \frac{I_{01}^D}{P_0 \cdot P}\}$ (the number of people, from country 0 to country 1, who are allowed in, divided by the number of those who try to get in). It is then possible to derive testable predictions for the impact of μ_1^0, μ_0, and μ_C on the emigration rate from country 0 to country 1:[24]

$$\frac{d(\frac{I_{01}}{P_0})}{d\mu_1^0} = \begin{cases} \frac{\phi(z)}{\sigma_v} > 0, \text{ if } \frac{I_{01}^S}{P_0} < \frac{I_{01}^D}{P_0}; \\ 0, \text{ if } \frac{I_{01}^S}{P_0} \geq \frac{I_{01}^D}{P_0} \end{cases} \tag{5}$$

$$\frac{d(\frac{I_{01}}{P_0})}{d\mu_h} = \begin{cases} -\frac{\phi(z)}{\sigma_v} < 0, \text{ if } \frac{I_{01}^S}{P_0} \leq \frac{I_{01}^D}{P_0}; \\ 0, \text{ if } \frac{I_{01}^S}{P_0} > \frac{I_{01}^D}{P_0} \end{cases} \tag{6}$$

[21] This assumption is consistent with the evidence that immigrants often arrive to a destination country with temporary tourist or student visas with the hope of being able to stay.

[22] In particular, $\sigma_v^2 = (\eta_{01}^2\sigma_1^2 + \sigma_0^2 + \sigma_C^2 - 2\eta_{01}\rho_{01}\sigma_0\sigma_1 - 2\eta_{01}\rho_{1C}\sigma_1\sigma_C + 2\rho_{0C}\sigma_0\sigma_C)$.

[23] The emigration rate we observe in the data, $\frac{I_{01}}{P_0}$, equals the actual number of immigrants from country 0 to country 1, divided by the population of country 0.

[24] The total differential of P equals:
$dP = \frac{\phi(z)}{\sigma_v} d(\eta_{01}\mu_1^0 - \mu_0 - \mu_C) + \phi(z)\ (\eta_{01}\mu_1^0 - \mu_0 - \mu_C)(-\frac{1}{\sigma_v^2})d\sigma_v$.

where $\phi(\cdot)$ is the density function of a standard normal and $h = 0, C$. According to (5) pull effects (namely, improvements in the mean income opportunities in the destination country) are positive and strongest when restrictions are not binding neither ex-ante nor ex-post, they are positive but smaller in size when the quota is binding ex-post but not ex-ante and, finally, they are equal to zero in a quantity-constrained world. A parallel interpretation explains the comparative-static results in (6), which describe push effects (changes of μ_0, that is mean income opportunities in the origin country) and the impact of mean migration costs (changes of μ_C), according to the immigration-policy regime.

Thus, according to this simple model, pull and push factors have either similar-sized effects (with opposite signs), when quotas are not binding, or they both have no (or a small) effect on emigration rates, when quotas are binding. In the empirical analysis I will not be able to control for whether migration quotas are binding for a country pair in a given year (since I do not have data on I_{01}^D). Therefore I will estimate an average effect across country pairs with different degrees of restrictiveness. However, I will be able to use information on changes in I_{01}^D: I should find that pull (push) effects are more positive (negative) than average, for a given destination country, if that country's migration policy becomes less restrictive.[25]

Focusing for simplicity on the region where immigration quotas are not binding, it is straightforward to derive predictions for the impact of second moments of the income distributions (of the origin country's population) at home and abroad respectively. In particular, assuming that $\sigma_C = 0$, we obtain the following expressions, where $k < 0$ (Borjas 1987):[26]

$$\frac{d(\frac{I_{01}}{P_0})}{d\sigma_1} = k \cdot (\mu_1^0 - \mu_0 - \mu_C) \cdot (\sigma_1 - \rho_{01}\sigma_0), \tag{7}$$

$$\frac{d(\frac{I_{01}}{P_0})}{d\sigma_0} = k \cdot (\mu_1^0 - \mu_0 - \mu_C) \cdot (\sigma_0 - \rho_{01}\sigma_1). \tag{8}$$

In my discussion I will assume that $(\mu_1^0 - \mu_0 - \mu_C) > 0$ so that, based on first-moments considerations, on average immigrants have an incentive to migrate. The results in (7) and (8) imply that, if $\frac{\sigma_0}{\sigma_1} < 1$ and ρ_{01} is sufficiently high ($\rho_{01} > \frac{\sigma_0}{\sigma_1}$), then $d\sigma_0 > 0$ or $d\sigma_1 < 0$ (i.e., an increase in the relative inequality $\frac{\sigma_0}{\sigma_1}$) will increase the emigration rate. Similarly, if $\frac{\sigma_0}{\sigma_1} > 1$ and ρ_{01} is sufficiently high ($\rho_{01} > \frac{\sigma_1}{\sigma_0}$), then $d\sigma_0 > 0$ or $d\sigma_1 < 0$ (i.e., an increase in the relative inequality $\frac{\sigma_0}{\sigma_1}$) will decrease the emigration rate.

3 Data

In this paper, I merge data from an international migration panel with macroeconomic and other information on the origin and destination countries of immigrant flows. Data on

[25] The reason is that, with higher I_{01}^D, the range of μ_1^0 (μ_0) for which the effect is strictly positive (negative) is wider (see Figure 2).

[26] Formulas (7) and (8) are based on the expression for dP in footnote 24 (given that immigration quotas are not binding, then $\frac{I_{01}}{P_0} = P$). If quotas are not binding ($\eta_{01} = 1$), assuming that $\sigma_C = 0$, then: $d\sigma_v = [\sigma_1^2 + \sigma_0^2 - 2\rho_{01}\sigma_0\sigma_1]^{-\frac{1}{2}}[(\sigma_1 - \rho_{01}\sigma_0)d\sigma_1 + (\sigma_0 - \rho_{01}\sigma_1)d\sigma_0 - \sigma_0\sigma_1 d\rho_{01}]$. Notice that, in formula (7) and (8), $k = \phi(z)(\sigma_1^2 + \sigma_0^2 - 2\rho_{01}\sigma_0\sigma_1)^{-\frac{1}{2}}(-\frac{1}{\sigma_v^2}) < 0$.

immigration comes from the International Migration Statistics (IMS) data set for OECD countries (OECD 1997), which provides information on bilateral immigrant flows based on the OECD's Continuous Reporting System on Migration (SOPEMI).[27] In particular, I use data on yearly immigrant inflows into fourteen OECD countries by country of origin, in the period 1980-1995.[28] Appendix 1, at the end of the paper, presents summary statistics on immigrant inflows by host and source country, averaged over the years 1980-1995. It shows that labor movements to the fourteen OECD countries are both South-North and North-North flows. The sample includes seventy-nine origin countries with per worker GDP levels ranging from approximately \$1,000 to \$55,000 (PPP-adjusted) on average in the period considered. In interpreting the numbers in Appendix 1, notice that the IMS data only covers legal immigration; population registers and residence and work permits are the main sources of these statistics.[29]

The quality of the IMS data is high even though the coverage is not complete. The data set is supposed to cover immigrant inflows into each of the fourteen destination countries from all over the world. However, the sum by country of origin of the IMS numbers is not equal to 100% of the total flow into each destination country. The percentage of the total immigrant inflow covered by the disaggregate data ranges between 45% (Belgium) and 84% (United States). Put differently, the data set includes zero flows in correspondence of some country pairs (immigrant inflows from Italy to the United States, for example): some of these observations correspond to truly zero flows, while others are likely to correspond to very small flows. If the latter observations are recorded as zeros in the disaggregate data set, there will be a discrepancy between total flows and the sum of flows by origin country. In the empirical analysis I will keep zero-flows observations in the data set. I will investigate the robustness of my results to excluding zero-flows observations and to using a Tobit model.

Summary statistics and data sources for the other regressors used in the empirical model are documented in Appendix 2. Data on macroeconomic variables comes from various sources: the 2001 World Development Indicators data set (World Bank 2001), the Penn World Tables (versions 5.6 and 6.1), and the World Bank's Global Development Network Growth Database, Macro Time Series (Easterly and Sewadeh 2002). Geographic and cultural information, such as on great-circle distance[30], land border, common language, and colonial ties, comes from Glick and Rose's (2002) data set on gravity-model variables. I also use statistics on the average number of schooling years in the total population of destination and origin countries (over age 15) from Barro and Lee's (2000) data set.[31] Data on

[27] Alvarez-Plata, Brücker, and Siliverstovs (2003) and Pedersen, Pytlikova, and Smith (2004) use different international-migration data sets: the former paper uses the Eurostat Labor Force Survey which covers all destination countries within the EU-15 over nine years; the latter paper uses a dataset constructed by the authors after contacting the statistical bureaus in 27 selected destination countries (this data set covers the years between 1990 and 2000).

[28] The OECD (1997) data can be accessed through SourceOECD. Unfortunately, data for the years after 1995 is not yet available.

[29] Although the migration data is not perfectly comparable across OECD countries (some countries in the OECD (1997) data set define immigrants based on country of birth, while others based on citizenship), it is reasonable to think that changes over time can be compared.

[30] Distance is calculated with the great circle formula using each capital city's latitude and longitude data.

[31] Since this panel only contains data at five-year intervals (in the period I consider, the years covered are 1980, 1985, 1990, 1995), I linearly extrapolate figures for the in-between years (by assigning one fifth of the

Gini coefficients of destination and origin countries, used to construct the origin country's relative inequality variable, comes from Deininger and Squire (1996) data set (I only used so called "high-quality observations")[32]. Finally, information on origin countries' share of young population comes from the United Nations.

Figure 1 shows that many destination countries in the sample are characterized by substantial volatility of immigrant inflows year after year. An important cause of variation over time in the number of immigrants to a given destination country is changes in that country's migration policy. For example, the United States graph in Figure 1 displays a peak around the year 1990. This is not surprising given that an amnesty law, the Immigration Reform and Control Act, was passed in 1986 and put in effect in the following years, with the bulk of the legalizations taking place in 1989-1991. The graph for Japan, on the other hand, displays a sudden decrease in the total immigrant inflow around the year 1982, which is when the Immigration Control and Refugee Recognition Act was passed. A separate Appendix to the paper documents the main characteristics of the migration policies of the destination countries in the sample and the timing (after 1980) of changes in their legislations (Mayda and Patel, 2004).[33] A data set of destination countries' migration-policy changes, between 1980 and 1995, was constructed on the basis of the information in this Appendix and used in the empirical analysis.[34]

4 Empirical model

According to the theoretical framework in Section 2, the estimating equation should be characterized by the emigration rate as the dependent variable and, among the explanatory variables, the mean wage of the origin country's population in, respectively, the origin and destination countries. As approximations for the latter two variables, I use the (log) level of per worker GDP, PPP-adjusted (constant 1996 international dollars) in the two countries.[35] Based on the theoretical model, I expect pull and push effects to be, respectively, positive and negative on average, if migration quotas are not binding, and both zero (or small) otherwise.

Another determinant of bilateral immigration flows implied by the model of Section 2 is the physical distance between the two locations, which affects migration costs C_i. The further away the two countries are, the higher the monetary travel costs for the initial move, as well as for visits back home. Remote destinations may also discourage migration because they

five-year change in the variable to each year).

[32] I linearly extrapolate data on Gini coefficients for the years in which it is not available, based on the values for other years for the same country.

[33] The Appendix (Mayda and Patel 2004) can be found at the author's Georgetown University website.

[34] In particular, the information in the Appendix (and in the background papers listed in the References) was used to identify: first, the timing of immigration-policy changes taking place in each destination country (the years in which migration policy laws were passed or enforced); second, the direction of the change in the case of substantial changes (loosening vs. tightening), based on a qualitative assessment of the laws (we mainly focused on aspects of migration policies related to the size of immigration flows, as opposed to, for example, issues of citizenship).

[35] Unfortunately, wage data cannot be used because wage income series are not available for all countries (especially origin ones) in the sample. Since per worker GDP is not a direct measure of the mean wage of the origin country's population at home and abroad, I run robustness checks at the end of the paper to check that it is a good proxy for it.

require longer travel time and thus higher foregone earnings. Another explanation as to why distance may negatively affect migration is that it is more costly to acquire information ex-ante about far-away countries (Greenwood 1997, Lucas 2000). Besides distance, I introduce additional variables that affect the level of migration costs C_i. A common land border is likely to encourage migration flows, since land travel is usually less expensive than air travel. Linguistic and cultural similarity are also likely to reduce the magnitude of migration costs, for example by improving the transferability of individual skill from one place to the other. Past colonial relationships should increase emigration rates, to the extent that they translate into similar institutions and stronger political ties between the two countries, thus decreasing the level of migration costs C_i.

In a cross-country analysis, such as in this paper, unobserved country-specific effects could result in biased estimates. For example, the estimate of the coefficient on the destination country's per worker GDP may be positive. Based on this result, it is not clear whether immigrants go to countries with higher wages or, alternatively, whether countries with higher wages have other characteristics that attract immigrants. Along the same lines, a negative coefficient on income at home leaves open the question of whether immigrants leave countries with lower wages or, alternatively, whether countries with lower wages have certain features that push immigrants to leave. To (partly) get around this problem, I exploit the panel structure of the data set and I introduce dummy variables for both destination and origin countries. This allows me to control for unobserved country-specific effects which are additive and time-invariant.[36] All the regressions also have robust standard errors clustered by country pair, to address heteroscedasticity and allow for correlation over time of country-pair observations. Notice that destination countries' fixed effects also allow me to control for features of their immigration policy which are time-invariant and common across origin countries. In order to capture the effect of *changes* in destination countries' migration policies, I introduce two interaction terms of an indicator variable of such changes with pull and push factors, respectively. According to the theory, if the migration policy of a destination country becomes less restrictive, the effect of pull (push) factors should turn more positive (negative).

Finally, I introduce the share of the origin country's population who is young (between 15 and 29 years old) as a demographic determinant of migration flows. Consider an extension of the basic model in Section 2 to a multi-period setting. In this set-up, the individual cares not only about current wage differentials net of moving costs, but about future ones too. As Clark, Hatton, and Williamson (2002) point out, this implies that a potential migrant from country 0 will have a bigger incentive to migrate the younger he is, as the present discounted value of net benefits will be higher the longer the remaining work life time is (for positive I_i in each year). We would then expect the share of the young population in the origin country to positively affect the emigration rate out of that country.

The basic empirical specification thus looks as follows:[37]

[36] In one robustness check, I control for country-pair fixed effects. In all the other regressions (based on pooled data), I include separate destination and origin countries' fixed effects.

[37] The basic empirical specification below also includes destination and origin countries' fixed effects, as explained above.

$$\frac{flow_{ijt}}{P_{it}} = \beta + \beta_0 pwgdp_{it-1} + \beta_1 pwgdp_{jt-1} + \beta_2 dist_{ij} + \beta_3 border_{ij} + \beta_4 comlang_{ij} + \beta_5 colony_{ij} + \\ + \beta_6 pwgdp_{it-1} \cdot immigpol_{jt} + \beta_7 pwgdp_{jt-1} \cdot immigpol_{jt} + \beta_8 youngpop_{it-1} + \varepsilon_{ijt} \quad (9)$$

where i is the origin country, j the destination country and t time. $\frac{flow_{ijt}}{P_{it}}$ is the emigration rate from i to j at time t ($flow_{ijt}$ is the inflow into country j from country i at time t, P_{it} is the population of the origin country at time t). *pwgdp* is the (log) per worker GDP, PPP-adjusted (constant 1996 international dollars) and *dist* measures the (log) great-circle distance between the two countries. The variable *border* equals one if the two countries in the pair share a land border. *comlang* and *colony* are two dummy variables equal to one, respectively, if a common language is spoken in the two locations, and for pairs of countries which were, at some point in the past, in a colonial relationship. The variable *immigpol* increases by one (decreases by one) if in that year the destination country's immigration policy became less (more) restrictive, zero otherwise.[38] Finally, *youngpop* is the share of the population in the origin country aged 15-29 years old. According to the model in Section 2, I expect that $\beta_0 \leq 0$, $\beta_1 \geq 0$, $\beta_2 \leq 0$, $\beta_3 \geq 0$, $\beta_4 \geq 0$, $\beta_5 \geq 0$, $\beta_6 < 0$, $\beta_7 > 0$, and $\beta_8 \geq 0$.

An econometric complication is the possibility of reverse causality and, more in general, of endogeneity in the time-series dimension of the analysis. For example, the theoretical model in Section 2 predicts that, *ceteris paribus*, better (worse) income opportunities in the destination (origin) country increase emigration rates. However, a positive β_1 (negative β_0) may just reflect causation in the opposite direction, that is the impact of immigrant flows on wages in host and source countries. After all, this channel is the main focus of analysis in many labour-economics papers (see Friedberg and Hunt (1995) for a survey of this literature).[39] More broadly, other time-variant third factors may drive contemporaneous wages and immigrant flows.

As for reverse causality, notice that it is likely to bias the estimates toward zero. The reason is that, if anything, immigrant inflows are likely to decrease wages in the destination country and outflows are likely to increase wages in the origin country. While the opposite signs are a theoretical possibility (for example, in the economic-geography literature, because of economies of scale), the empirical evidence in the labor-economics literature is that immigrant inflows have a negative or zero impact on the destination country's wages (Friedberg and Hunt 1995, Borjas 2003) and that immigrant outflows have a positive impact on the origin country's wages (Mishra 2003).

Although reverse causality may not be an issue, it is still important to address other sources of endogeneity, in the following two ways. First of all, in the basic specification, I relate *current* emigration rates to *lagged* values of (log) per worker GDP, at home and abroad (I use lagged values also for all the other time-varying regressors). While it is unrealistic to claim that wages at home and abroad are strictly exogenous, it is plausible to assume that they are predetermined, in the sense that immigrant inflows - and third factors in the error

[38] In other words, a change in policy is modelled as leading to a lasting effect (i.e., in the year when the policy change occurred and in the following years).

[39] At the same time, given that in this paper the dependent variable is immigration *flows* (as opposed to *stocks*), reverse causality may be less of an issue.

term - can only affect contemporaneous and future wages.[40] As a robustness check, I also use instrumental-variables estimation with countries' terms of trade as an instrument for PPP-adjusted income levels in destination and origin countries. Papers in the literature where shocks to terms of trade are used as instruments for growth rates of income are, for example, Pritchett and Summers (1996) and Easterly, Kremer, Pritchett and Summers (1993).

5 Empirical results

Table 1 presents the results from estimation of equation (9) controlling for destination and origin countries' fixed effects. The estimates show a systematic pattern, broadly consistent with the theoretical predictions of the international migration model. The analysis also generates empirical puzzles.

First, the emigration rate is positively related to the destination country's (log) per worker GDP.[41] According to the estimate in regression (1), a ten percent increase in the level of per worker GDP in the destination country increases emigration by 2.5 emigrants per 100,000 individuals of the origin country's population (significant at the 5% level). In other words, a 10% increase in the host country's per worker GDP implies a 19% increase in the emigration rate (as the mean of the dependent variable is, in regression (1), 13 emigrants per 100,000 individuals). This result would suggest that migration quotas are not binding on average across destination countries. However, the impact on the emigration rate of a change in the income opportunities at home is not consistent with this interpretation. Push effects are estimated to be insignificantly different from zero in Table 1. One possibility is that, in practice, migration quotas are not binding, but push factors are zero due to the effect of poverty constraints in the origin country. I will investigate this hypothesis in Table 2.

In regressions (1)-(3), Table 1, I also explore the role played by geographic (*log distance* and *land border*), cultural (*common language* and *colony*), and demographic (*share of young population* (*origin*)) determinants, respectively. The picture that emerges from my results is one in which geography and demographics are the most important among this set of drivers of migration flows. According to the estimate in column (1), doubling the great-circle distance between the source and host country decreases the number of emigrants by 41 per 100,000 individuals in the origin country (significant at the 1% level). On the other hand, a common land border does not appear to play a significant role. The impact of a common language, though of the right sign, is not statistically significant and, surprisingly, past colonial relationships do not appear to affect migration rates.[42] Finally, the share of the origin country's population who is young has a positive and significant impact on emigration rates. A ten percentage point increase in the origin country's 15-29 years old population

[40] Strict exogeneity of an explanatory variable implies $E[X_{it}\varepsilon_{is}] = 0$, for $\forall s,\ t$, while predeterminacy implies $E[X_{it}\varepsilon_{is}] = 0$, for $\forall s > t$. In one of the following specifications, I also control for lagged values of the emigration rate, since if the emigration rate is autocorrelated, predeterminacy of the regressors does not guarantee consistency of the estimates.

[41] I constrain the sample of observations to be the same in the *pooled* regressions of this table and of the other tables (whenever data availability makes it possible).

[42] This statement is true whether *common language* and *colony* are entered in the regression together or one at a time.

raises the emigration rate by 24 emigrants per 100,000 individuals (regression (3)). These results are confirmed in column (4) where, out of all the geographic, cultural and demographic determinants, I only include the ones which are significant based on regressions (1)-(3), that is *log distance* and *share of young population* (*origin*).

In the next regression (column (5)) I only exploit the variation over time within country pairs, by introducing fixed effects for each combination of origin and destination countries.[43] These country-pairs dummy variables allow me to control for time-invariant features of the destination country's immigration policy which are specific for each origin country. The results from this specification confirm that push and pull factors have an asymmetric effect in terms of magnitudes and significance levels.[44]

The framework used in regressions (1)-(5) to study migration flows is related to the gravity model of trade, which is employed to analyze bilateral *trade* flows across countries.[45] As a matter of fact, I use several variables that appear frequently in the trade gravity literature (*log distance*, *land border*, *common language*, and *colony*). The specifications in the following three columns ((6)-(8)) use the same regressors as in regression (4) but are more closely related to trade gravity-model regressions, which are usually estimated in a cross section. That is, in regressions (6)-(8) I only exploit the cross-country variation in the data by estimating the model year by year (I focus on three years: 1985, 1990, and 1995)[46] Due to the low number of observations in each year, I do not control for country-specific fixed effects, which could explain the difference in magnitudes of the effects relative to previous regressions. However, the coefficients are still qualitatively consistent with the panel-data estimates.

Next, I investigate the interaction between changes in destination countries' migration policies and, respectively, pull and push factors (column (9), Table 1). Consistent with the theoretical predictions, positive pull factors are bigger than average for a destination country whose migration policy becomes less restrictive. Setting aside the average effect, push factors turn negative and significant once migration restrictions are relaxed. The opposite is true when policy becomes more protectionist. In the same regression I also add the interaction of the indicator variable of changes in destination countries' migration policy with, respectively, *log distance* and *share of young population* (*origin*). I find that the effect of the latter two variables is more pronounced (more negative and more positive, respectively) when a host country's immigration laws turn less restrictive. The opposite is true when policy becomes more protectionist. These results do no change when I include the main effect of

[43] Therefore I do not include the regressors *log distance*, *land border*, *common language* and *colony* since they are constant within country pairs and, therefore, would be perfectly collinear with the dummy variables.

[44] If country pairs differ in terms of out-migration and return migration rates, net migration flows can be very different from gross flows. Since out-migration and return migration are likely to characterize specific country pairs, they are accounted for by including country-pair fixed effects.

[45] There exists a gravity model of immigration, developed in the geography literature and sometimes used in economics papers. The empirical specification I use, suggested by economic theory, differs in part from the standard equation estimated by geographers, which looks as follows (Gallup 1997): $flow_{ij} \propto \frac{P_i P_j}{dist_{ij}^2}$. That is, there is still a contrast between economic and gravity explanations of immigrant flows (Helliwell 1997).

[46] I also estimate the model in 1981-1985, in 1986-1990 and in 1991-1995 (results not shown) and get very similar estimates (that is, I confirm the asymmetry between pull and push factors). In particular, in these regressions, I relate average emigration rates in each subperiod to the average income opportunities at home and abroad in the previous five-year interval (plus time-invariant variables).

immigration policy changes (regression (10)), which has an insignificant impact. Regression (9) represents the preferred specification of the model. It shows that migration restrictions matter by mitigating effects on the supply side of the model (pull and push factors, geography and demographics).

In Table 2, I analyze economic determinants more in detail. First, I investigate the impact of the second moments of the income distributions in the origin and destination countries. According to the theory (formulas (7) and (8)), given low values of the origin country's relative inequality ($\frac{\sigma_0}{\sigma_1}$), if $\frac{\sigma_0}{\sigma_1}$ increases, the emigration rate will increase, while given high values of $\frac{\sigma_0}{\sigma_1}$, if $\frac{\sigma_0}{\sigma_1}$ increases, the emigration rate will decrease.[47] The intuition for these results is straightforward. If income inequality in the origin country is lower than in the destination country ($\frac{\sigma_0}{\sigma_1} < 1$), there is positive selection of immigrants from country 0 to country 1: migrants are selected from the upper tail of the income distribution at home and end up in the upper tail of the income distribution abroad (in both cases, the relevant distribution is the origin country's population one). For example, consider potential migrants from Portugal to the United States. Given that income inequality is lower in Portugal than in the U.S., among Portuguese workers it is the better-off who have an incentive to migrate while those at the very low tail of the income curve have an incentive to stay. The reason is that the probability of both very high and very low incomes is higher in the U.S. than in Portugal. An increase in income inequality in Portugal will make the marginal individual (who is in the lower tail of the income distribution) relatively worse-off at home and will increase her incentive to leave. Similarly, if income is more dispersed at home than abroad ($\frac{\sigma_0}{\sigma_1} > 1$), then there is negative selection of immigrants from country 0 to country 1: migrants are selected from the lower tail of the income distribution at home and end up in the lower tail of the income distribution abroad. An example of this situation is migration from Brazil to the U.S., given that income inequality in the latter is lower than in the former.[48] An increase in income inequality in Brazil will lower the emigration rate because those who were not migrating beforehand, the better-off, will have even less incentive to do so afterwards. In order to test these predictions, I introduce in the estimating equation a measure of the origin country's relative inequality ($\frac{\sigma_0}{\sigma_1}$) both in linear and quadratic forms. As expected, I find that the coefficient on the linear term is positive and on the quadratic term is negative (both significant at the 5% level), which is consistent with Borjas (1987) selection model (regressions (1)-(2)[49], Table 2).[50]

---

[47] I assume that ρ_{01} is sufficiently high ($\rho_{01} > \max\{\frac{\sigma_0}{\sigma_1}, \frac{\sigma_1}{\sigma_0}\}$). The motivation for this assumption is explained in Borjas (1987): "It seems plausible to argue that for non-Communist countries, ρ_{01} is likely to be positive and large. After all, profit-maximizing employers are likely to value the same factors in any market economy" (p.534). I also assume that $(\mu_1^0 - \mu_0 - \mu_C) > 0$ so that, based on first-moments considerations, on average immigrants have an incentive to migrate. The motivation for the last assumption is that the data set mostly includes migration flows from lower to higher average-income countries: the average difference in per capita GDP levels of destination and origin countries is positive and substantial (approximately \$20,600). I also add a robustness check (regression (2), Table 2) where I only include observations characterized by a positive difference between the per capita GDP levels of destination and origin countries in any given year.

[48] The Gini coefficient for Portugal was 36.76 in 1990, while in the U.S. it was 37.8. The Gini coefficient for Brazil was 61.76 in 1985, while in the U.S. it was 37.26 (Deininger and Squire 1996).

[49] See footnote 47 for a discussion of regression (2), Table 2.

[50] I evaluate the effect of relative inequality over the relevant range of values. Based on the coefficient estimates in column (1), Table 2, the threshold value of relative inequality is approximately equal to 2.73805:

The remaining specifications in Table 2 investigate empirically a few extensions of the theoretical framework of Section 2. First, it is possible to incorporate poverty constraints in the model, due to fixed costs of migration and credit market imperfections in the origin country. As Yang (2003) shows, these assumptions imply that the effect on emigration rates of income opportunities at home is non-monotonic, positive at very low levels of income and negative for higher levels. Accordingly, I extend the empirical model previously specified by introducing both a linear and a quadratic term in per worker GDP of the origin country. I find only weak evidence of poverty constraints in regression (3). The sign of the coefficients is consistent with the theory but the low significance level of the estimates prevents me from reading too much support into them.[51] This result thus leaves unanswered the question of why push and pull effects are different in size and, indirectly, lends support to the alternative hypothesis of binding (and endogenous) migration quotas.

Next, the theoretical model can be modified by taking into account uncertainty in finding a job in each place. This extension suggests using the unemployment rate (which is approximately equal to one minus the probability of finding a job) as a regressor in the estimating equation. My results in column (4) confirm the asymmetry between push and pull factors, this time in terms of unemployment rates. In an additional extension (column (5)), I test whether workers choose among multiple destination countries. In the theoretical model, the choice is between the origin country and one particular destination country. In practice, however, potential migrants are likely to compare mean income opportunities in their origin country to those in the destination country considered *and in any other host country.* For each pair of source and host economies, I construct and control for a *multilateral pull* term which is an average of per worker GDP levels of all the other destination countries in the sample, each weighted by the inverse of distance from the origin country. Regression (5) shows that third-country effects shape bilateral migration flows as expected, given that the coefficient on the *multilateral pull* term is indeed negative and significant (at the 5% level).[52]

To conclude, I investigate the role of past migration flows to the destination country from the same origin country. Lagged emigration rates capture the impact of network effects, which are likely to reduce the cost C_i of migration. The introduction of the lagged emigration rate among the explanatory variables makes the model a dynamic one. I use Arellano and Bond's GMM estimator to deal with the incidental parameter problem that arises with fixed-effects estimation of such a dynamic equation.[53] Emigration rates show considerable inertia in regression (6), where the coefficient on the lagged emigration rate is 0.66 (significant

---

if $\frac{\sigma_0}{\sigma_1}$ is below this value (which is the case in my sample, based on the summary statistics in Appendix 2), an increase in $\frac{\sigma_0}{\sigma_1}$ raises the emigration rate which is consistent with positive selection taking place.

[51] In contrast with my results – which are only marginally significant – Hatton and Williamson (2003) and Pedersen, Pytlikova and Smith (2004) find evidence of an inverted U-shaped effect on emigration of the origin country's economic conditions.

[52] The multilateral pull term places migrants' decision to move in a multi-country framework. It is inspired by the multilateral trade resistance term in Anderson and van Wincoop (2003) (even though mine is an atheoretical measure).

[53] The Arellano and Bond estimator transforms into a difference the initial equation to remove the country-pair fixed effect and produces an equation that can be estimated with instrumental variables using a generalized method-of-moments estimator. The instruments include the lagged values of the dependent variable starting from t-4-2 (since the regression includes, as regressors, the emigration rate lagged by one, two, three and four years).

at the 1% level).[54] However, outside the model of Section 2 – which assumes exogenous migration quotas – it is unclear how to interpret this autocorrelation. While it is consistent with network effects on the supply side, it could also be driven by factors working on the demand side. In particular, through the latter channel, past migration flows can influence the emigration rate in two different ways: through family-reunification immigration policies and through political-economy factors (see, for example, Goldin (1994) and Ortega (2005), where the votes of naturalized immigrants affect immigration-policy outcomes).

6 Robustness checks

In Table 3 I run a few robustness checks of my previous results. In the first regression, I use (within-country variation in) the terms of trade of destination and origin countries, respectively, to instrument for (within-country variation in) their level of per worker GDP. Since terms of trade are defined as the price of a country's exports divided by the price of its imports, they represent how much a country can import per unit of exported good. Thus, they are correlated with the average real income of a country because they affect its purchasing power vis a vis goods produced by the rest of the world. In the first stage, the impact of the terms of trade of destination countries (origin countries) on their per worker GDP is positive and significant at the 1% level. In addition, the F value for the excluded instruments (terms of trade) is equal to 30.47 and 70.36 in the first stage of per worker GDP of destination and origin countries, respectively. Finally, if the assumption of small open economies holds[55], terms of trade are unlikely to affect emigration rates directly or to be correlated with other country-level characteristics that have an impact on migration patterns (exclusion restriction). Regression (1) shows that pull and push coefficients are robust to endogeneity issues as the instrumental variable estimates are consistent with my earlier results. In particular, this regression confirms the asymmetry between push and pull factors.

My second robustness check investigates whether per worker GDP (PPP-adjusted) of origin and destination countries is a good proxy for mean income opportunities of migrant workers at home and abroad. Per worker GDP is not a direct measure of wages of a potential migrant, since it depends on rates of return to both capital and labor and on endowments of each factor. For example, a higher per worker GDP in the destination country does not necessarily mean better income opportunities on average for an immigrant worker, since it could be due to a higher capital-labor ratio or to a more skilled labor force in the destination country's population. To address this concern, I run a robustness check where I control for the mean skill level and per worker endowment of capital in each country (columns (2) and (3))[56]. I first control for the average schooling level in both countries in specification (2).

[54] Regression (6) includes, as regressors, the emigration rate lagged by one, two, three and four years (the coefficients on the latter three lags are not shown in the table). The reason is that, only by introducing all these lags, I don't reject the null of zero autocovariance in residuals of order 2 (which is one of the requirements of the Arellano and Bond estimator).

[55] There is debate about how large a country must be in order to have an impact on its terms of trade. Irwin (2002) argues that the United States, for example, should be considered as small in almost all markets, while Feenstra (2004) has the opposite view.

[56] Since capital is assumed to be internationally mobile, there are no international differences in rates of

I still estimate pull effects which are positive and significant (at the 1% level). The results on push effects are the same as in previous estimates as well. In line with the theoretical predictions, the average skill level in the population of the destination (origin) country has a negative (positive) impact on the emigration rate. In column (3) I control for per worker endowments of both skill and capital and find that their coefficients are of the right sign. In addition, my prior findings on pull and push factors are robust. Finally, the last two regressions in Table 3 (columns (4)-(5)) test how robust the results are – in particular, in terms of the asymmetry between pull and push factors – to excluding zero-flows observations and to using a Tobit specification. These estimates are again consistent with the picture based on OLS and IV regressions.

7 Conclusions

In this paper, I empirically investigate the determinants of international bilateral migration flows. This analysis both delivers estimates consistent with the predictions of the international migration model and generates empirical puzzles.

In particular, I find evidence that pull factors, that is improvements in the mean income opportunities in the destination country, significantly increase the size of emigration rates. This result is very robust to changes in the specification of the empirical model. On the other hand, the sign of the impact of push factors - that is, declining levels of per worker GDP in the origin country - is seldom negative and, when it is, the size of the effect is smaller than for pull factors and is almost always insignificant. Therefore the evidence uncovered by the estimates is mixed in terms of the migration-policy regime that characterizes, on average, the destination countries in the sample: Push effects suggest that migration quotas are more binding than pull effects do. A possible explanation of the asymmetry between push and pull factors is the role played by the demand side of the model, that is destination countries' migration policies. While the theoretical framework of Section 2 assumes that migration quotas are exogenous, in practice they are not. Indeed migration policies can be thought of as the outcome of a political-economy model in which voters' attitudes towards immigrants, interest-groups pressure, policy-makers preferences and the institutional structure of government interact with each other and give rise to a final immigration-policy outcome (Rodrik 1995, Facchini and Willmann 2005, Mayda 2006).[57] Binding and endogenous migration quotas can explain the asymmetric effect I estimate for pull and push factors. While I do not investigate the endogenous determination of migration policy, I find evidence consistent with the constraining role played by migration policies. In the empirical analysis, I interact an indicator variable of changes in destination countries' migration policies with pull and push factors, respectively. I find that pull effects become more positive and push effects turn negative in those years when a host country's immigration laws become less restrictive.

return to capital.

[57] In a political-economy framework, the immigration quotas are likely to depend on the capital-labor ratio of the median voter (see Benhabib 1996), on the size of past immigration flows from the same origin country, both because of family-reunification policies and because of pro-immigration votes of naturalized immigrants (Ortega 2005), on the extent of political organization of various interest groups (Grossman and Helpman 1994, Facchini and Willmann 2005) and on the business cycle.

Among the variables affecting the costs of migration, distance appears to be the most important one. Its effect is negative, significant and quite steady across specifications. Demographics, in particular the share of the origin country's population who is young, represent a significant determinant of emigration rates as well. I find that the effect of both variables is more pronounced in those years when a host country's immigration laws become less restrictive. In sum, my results suggest that migration quotas matter: They mitigate supply-side effects, that is pull and push factors, geography and demographics.

The investigation of the determinants of international migration leads to other interesting research questions. As already pointed out, the framework used in this paper to study migration flows is related to the gravity model of trade, which is used to analyze bilateral *trade* flows across countries. A common framework of empirical analysis for trade and migration makes it possible to combine the study of these two dimensions of international integration. In addition, while this paper looks at the determinants of international migration, it provides a framework to analyze the impact of migration, on source and host economies - on their standards of living, for example. The analysis of the impact of migration is usually affected by problems of reverse causality and endogeneity. Since my paper helps identify the determinants of immigrant flows - some of which are possibly exogenous - it is possible to use its results to construct a *first stage* (as in the trade gravity-model literature – see Frankel and Romer (1999), for example).

To conclude, by taking advantage of both the time-series and cross-country variation in an annual panel data set, this paper makes progress in explaining the determinants of international migration flows and in providing a framework for future analyses of their impact.

References

Adams, R. H. J. (1993). The economic and demographic determinants of international migration in rural Egypt. *Journal of Development Studies*, 30(1):146–167.

Alvarez-Plata, P., Brucker, H., and Siliverstovs, B. (2003). Potential migration from Central and Eastern Europe into the EU-15 - An update. Report for the European Commission, DG Employment and Social Affairs.

Anderson, J. E. and Wincoop, E. V. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93:170–192.

Barro, R. and Lee, J. (2000). International data on educational attainment. Data Set.

Benhabib, J. (1996). On the political economy of immigration. *European Economic Review*, 40:1737–1743.

Borjas, G. and Bratsberg, B. (1996). Who leaves? The outmigration of the foreign-born. *Review of Economics and Statistics*, 78(1):165–176.

Borjas, G. J. (1987). Self selection and the earnings of immigrants. *American Economic Review*, 77:531–553.

Borjas, G. J. (1994). The economics of immigration. *Journal of Economic Literature*, pages 1667–1717.

Borjas, G. J. (1995). The economic benefits from immigration. *Journal of Economic Perspectives*, 9(2):3–22.

Borjas, G. J. (1999a). The economic analysis of immigration. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, chapter 28, pages 1697–1760. North-Holland Elsevier Science, The Netherlands.

Borjas, G. J. (1999b). *Heaven's Door: Immigration Policy and the American Economy.* Princeton University Press, Princeton, N.J.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics*, 118(4):1335–1374. Harvard University.

Brücker, H., Siliverstovs, B., and Trübswetter, P. (2003). International migration to Germany: Estimation of a time-series model and inference in panel cointegration. DIW Discussion paper No. 391.

Chiswick, B. R. and Hatton, T. J. (2002). International migration and the integration of labour markets. In *Globalization in Historical Perspective.* The University of Chicago Press, Chicago, IL. Forthcoming.

Clark, X., Hatton, T. J., and Williamson, J. G. (2002). Explaining US immigration, 1971-1998. National Bureau of Economic Research Working Paper No. 8998. Forthcoming in Review of Economics and Statistics.

Coppel, J., Dumont, J.-C., and Visco, I. (2001). Trends in immigration and economic consequences. OECD Economics Department Working Papers No. 284.

Davies, J. B. and Wooton, I. (1992). Income inequality and international migration. *The Economic Journal*, 102(413):789–802.

Davis, D. R. and Weinstein, D. E. (2002). Technological superiority and the losses from migration. National Bureau of Economic Research Working Paper No. 8971.

Easterly, W., Kremer, M., Pritchett, L., and Summers, L. H. (1993). Good policy or good luck? Country growth performance and temporary shocks. *Journal of Monetary Economics*, 32:459–483.

Easterly, W. and Sewadeh, M. (2002). Global development network growth database, Macro time series. World Bank. Data Set.

Facchini, G. and Willmann, G. (2005). The political economy of international factor mobility. *Journal of International Economics*, 67(1):201–219.

Faini, R. (2002). Discussion of "International migration and the integration of labor markets" by B. Chiswick and T. Hatton. In *Globalization in Historical Perspective.* The University of Chicago Press, Chicago, IL. Forthcoming.

Faini, R. and Venturini, A. (1994). Trade, aid and migrations. Some basic policy issues. *European Economic Review*, 37:435–442.

Feenstra, R. (2004). *Advanced International Trade: Theory and Evidence.* Princeton University Press, Princeton.

Frankel, J. A. and Romer, D. (1999). Does trade cause growth? *The American Economic Review*, 89(3):379–399.

Freeman, G. (1992). Migration policy and politics in the receiving states. *International Migration Review*, 26(4):1144–1167.

Freeman, G. (1995). Modes of immigration politics in liberal democratic states. *International Migration Review*, 29(4):881–902.

Friedberg, R. and Hunt, J. (1995). The impact of immigrants on host country wages, employment, and growth. *Journal of Economic Perspectives*, 9(2):23–44.

Glick, R. and Rose, A. (2002). Does a currency union affect trade? The time-series evidence. *European Economic Review*, 46(6):1125–1151.

Goldin, C. (1994). The political economy of immigration restriction in the United States, 1890 to 1921. In Goldin, C. and Libecap, G., editors, *The Regulated Economy: A Historical Approach to Political Economy*, pages 223–257. University of Chicago Press, Chicago, IL.

Greenwood, M. J. (1997). Internal migration in developed countries. In Rosenzweig, M. and Stark, O., editors, *Handbook of Population and Family Economics, Vol. 1B.* North-Holland, Amsterdam.

Hanson, G. and Woodruff, C. (2004). Emigration and educational attainment in Mexico.

Hanson, G. H. and Spilimbergo, A. (2001). Political economy, sectoral shocks, and border enforcement. *Canadian Journal of Economics*, 34(3):612–638.

Hatton, T. J. (2004). Emigration from the UK, 1870-1913 and 1950-1998. *European Review of Economic History*, 8:149–171.

Hatton, T. J. (2005). Explaining trends in UK immigration. *Journal of Population Economics*, 18:719–740.

Hatton, T. J. and Williamson, J. G. (2003). Demographic and economic pressure on emigration out of Africa. *Scandinavian Journal of Economics*, 105:465–482.

Hatton, T. J. and Williamson, J. G. (2005). What fundamentals drive world migration? In Borjas, G. and Crisp, J., editors, *Poverty, International Migration and Asylum*. Palgrave-Macmillan.

Helliwell, J. F. (1997). National borders, trade and migration. National Bureau of Economic Research Working Paper No. 6027.

Helliwell, J. F. (1998). *How Much Do National Borders Matter?*, chapter 5, pages 79–91. Brookings Institution Press.

Heston, A., Nuxoll, D. A., Summers, R., and Aten, B. (1995). Penn World Table version 5.6a. Data Set.

Heston, A., Summers, R., and Aten, B. (2002). Penn World Table version 6.1. Center for International Comparisons at the University of Pennsylvania (CICUP). Data Set.

Hoddinott, J. (1994). A model of migration and remittances applied to Western Kenya. *Oxford Economic Papers*, 46:459–476.

Hunt, J. (2006). Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association*, 4(5):1014–1037.

ILO (1998). International labour migration database. International Labor Organization.

Irwin, D. A. (2002). *Free Trade Under Fire*. Princeton University Press, Princeton.

Jenks, R. (1992). *Immigration and Nationality Policies of Leading Migration Nations.* Center for Immigration Studies, Washington, DC.

Joppke, L. (1998). Why liberal states accept unwanted immigration. *World Politics*, 50:266–293.

Karemera, D., Oguledo, V. I., and Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied Economics*, 32(13):1745–1755.

Lopez, R. and Schiff, M. (1998). Migration and the skill composition of the labor force: The impact of trade liberalization in LDCs. *Canadian Journal of Economics*, 31(2):318–336.

Lucas, R. E. (2000). The effects of proximity and transportation on developing country population migrations. Boston University.

Martin, P. (2004). Copenhagen challenge paper on population and migration.

Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *Review of Economics and Statistics*, 88(3):510–530.

Mayda, A. M. and Patel, K. (2004). OECD countries migration policies changes. Appendix to International migration: A panel data analysis of the determinants of bilateral flows, by Anna Maria Mayda. http://www.georgetown.edu/faculty/amm223/.

Mishra, P. (2007). Emigration and wages in source countries: Evidence from Mexico. *Journal of Development Economics*, 82(1):180–199.

Money, J. (1997). No vacancy: The political geography of immigration control in advanced industrial countries. *International Organization*, 51:685–720.

Nations, U. (2004). *World Economic and Social Survey 2004. International Migration.* United Nations, New York.

OECD (1997). International migration statistics for OECD countries. Data set.

OECD (2001). The employment of foreigners: Outlook and issues in OECD countries. In *OECD Employment Outlook*, chapter 5. OECD, Paris.

Ortega, F. (2005). Immigration policy and skill upgrading. *Journal of Public Economics*, 89:1841–1863.

Pedersen, P. J., Pytlikova, M., and Smith, N. (2004). Selection or network effects? Migration flows into 27 OECD countries, 1990-2000. IZA Discussion Paper 1104.

Pedersen, P. J., Pytlikova, M., and Smith, N. (2006). Migration into OECD countries 1990-2000. In Parson and Smeeding, editors, *Immigration and the Transformation of Europe.* Cambridge University Press, Cambridge.

Pritchett, L. (2003). The future of migration. *Yale Global Online*.

Pritchett, L. and Summers, L. H. (1996). Wealthier is healthier. *The Journal of Human Resources*, 31(4):841–868.

Ravenstein, E. (1885). The laws of migration. *Proceedings of the Royal Statistical Society*, XLVII(2):167–235.

Rodrik, D. (1995). Political economy of trade policy. In Grossman, G. and Rogoff, K., editors, *Handbook of International Economics, Vol.3*, chapter 28, pages 1457–1494. North-Holland Elsevier Science, The Netherlands.

Rodrik, D. (2002). Comments at the conference on "Immigration policy and the welfare state". In Boeri, T., Hanson, G. H., and McCormick, B., editors, *Immigration Policy and the Welfare System.* Oxford University Press.

SOPEMI (1997). *Trends in International Migration. Annual Report 1996.* OECD, Paris.

Trefler, D. (1993). International factor price differences. *Journal of Political Economy*, 101(6):961–987.

Yang, D. (2003). Financing constraints, economic shocks, and international labor migration: Understanding the departure and return of philippine overseas workers. Harvard University, dissertation chapter.

Yang, P. Q. (1995). *Post-1965 Immigration to the United States.* Praeger, Westport, Connecticut.

Table 1. Determinants of bilateral immigrant flows

Equation	1	2	3	4	5	6	7	8	9	10
Dependent variable	Emigration rate									
log per worker gdp (destination)	24.62	24.79	29.41	29.34	33.01	52.05	167.41	103.07	17.35	20.66
	11.30*	11.27*	11.48*	11.53*	12.55**	23.09*	57.55**	40.79*	8.15*	9.40*
log per worker gdp (origin)	-0.77	-1.03	3.32	3.94	-9.04	-2.4	-2.98	-1.44	7.63	7.45
	7.23	7.09	8.02	8.22	5.63	2.07	3.19	1.65	8.71	8.73
log distance	-41.01	-40.65	-40.66	-37.94		-9.61	-20.63	-10.94	-41.85	-41.84
	9.50**	9.08**	9.08**	8.00**		3.21**	6.18**	2.57**	8.41**	8.41**
land border	-28.16	-36.97	-36.95							
	19.67	23.23	23.28							
common language		22.05	22.03							
		15.87	15.87							
colony		3.03	2.89							
		16.89	16.93							
share of young population (origin)			242.36	248.25	165.76	292.87	521.77	155.71	281.48	283.68
			110.23*	112.35*	88.77+	118.63*	177.22**	60.80*	118.34*	116.99*
per worker gdp (destination)*immig policy change									7.56	17.17
									2.04**	5.84**
per worker gdp (origin)*immig policy change									-3.37	-3.2
									1.37*	1.44*
log distance*immig policy change									-10.2	-10.18
									2.50**	2.48**
share of young population (origin)*immig policy change									144.47	149.85
									48.43**	48.47**
immig policy change										-106.51
										69.14
number of observations	8010	8010	8010	8010	8010	551	606	650	8010	8010
R-squared	0.24	0.25	0.25	0.24	0.85	0.04	0.07	0.06	0.27	0.27

OLS estimates. Destination and origin countries' dummy variables are included in specifications (1)-(4) and (9)-(10). Regression (5) includes country-pair fixed effects. Regressions (6), (7), (8) are for the years 1985, 1990, 1995, respectively, and have no fixed effects. Standard errors, clustered by country pairs, are presented under each estimated coefficient. + significant at 10%; * significant at 5%; ** significant at 1%. Constant not shown. See Appendix 2 for data sources.

The *emigration rate* (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) gives the number of incoming immigrants per 100,000 individuals in the origin country's population. *per worker gdp* is the level of per worker GDP, PPP-adjusted (constant 1996 international dollars), lagged by one year. *distance* is the great-circle distance. *Land border* equals one if the destination and origin countries share a land border.

common language equals one if a common language is spoken in both host and origin countries. *colony* is a dummy variable for pairs of countries ever in a colonial relationship. *share of young population (origin)* is the share of the population in the origin country aged 15-29, lagged by one year. *immig policy change* increases by one if in that year the host country's immigration policy became less restrictive, decreases by one if it became more restrictive, zero if there was no change.

Table 2. Panel data regressions: Economic determinants more in detail

Equation	1	2	3	4	5	6
Dependent variable	**Emigration rate**					
log per worker gdp (destination)	38.46	50.43	31.92	2.37	38.22	37.78
	17.97*	20.78*	12.15**	16.83	13.31**	12.31**
log per worker gdp (origin)	7.44	6.55	83.12	1.66	5.23	1.68
	17.12	17.62	49.38+	12.09	8.21	7.15
square of log per worker gdp (origin)			-4.31			
			2.81			
origin country's relative inequality	77.35	82.41				
	39.05*	44.32+				
square of relative inequality	-28.25	-29.44				
	13.17*	14.51*				
unemployment rate (destination)				-1.36		
				0.45**		
unemployment rate (origin)				0.62		
				0.76		
multilateral pull					-6.41	
					3.18*	
emigration rate(t-1)						0.66
						0.02**
log distance			-36.85	-30.42	-36.37	
			7.57**	8.26**	7.43**	
common language			18.85	19.2	18.93	
			11.84	11.57+	11.85	
share of young population (origin)			225.31	57.12	239.64	
			110.81*	87.21	109.92*	
constant	-512.82	-631.03	-433.39	191.73	-159.84	-0.59
	181.56**	223.47**	271.59	136.56	131.71	0.24*
number of observations	4028	3350	8010	5010	8010	6429
R-squared	0.18	0.18	0.24	0.23	0.24	

OLS estimates, except for in regression (6) (see below). Destination and origin countries' dummy variables are included in each specification (except in regression (6)). Standard errors, clustered by country pairs, are presented under each estimated coefficient. + significant at 10%; * significant at 5%; ** significant at 1%. See the end of Table 1 for definitions of the main variables used.

The *emigration rate* (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) gives the number of incoming immigrants per 100,000 individuals in the origin country's population.

multilateral pull gives, for each destination/origin country pair, the average of (log per worker gdp (destination)-log distance) over all the other destination countries. *origin country's relative inequality* gives a measure of the inequality in the origin country relative to the destination country (it equals the gini coefficient in the origin country divided by the gini coefficient in the destination country).

In regression (2), I only include observations characterized by a positive difference between the per capita GDP levels of destination and origin countries in any given year. In equation (6), I include as regressors the emigration rate lagged by one, two, three and four years (the coefficients on the latter three lags are not shown). The reason is that, only by introducing all these lags, I don't reject the null of zero autocovariance in residuals of order 2 (which is one of the requirements of the Arellano and Bond estimator). Column (6): Arellano-Bond test that average autocovariance in residuals of order 1 is 0: H0: no autocorrelation z = -55.05 Pr > z = 0.0000. Column (6): Arellano-Bond test that average autocovariance in residuals of order 2 is 0: H0: no autocorrelation z = -0.35 Pr > z = 0.7269. See Appendix 2 for data sources.

Table 3. Panel data regressions: Robustness Checks

Equation	1	2	3	4	5
Method	Instrumental Variables Estimation	OLS	OLS	OLS	TOBIT
Dependent variable	Emigration rate				
log per worker gdp (destination)	40.91	37.41	52.96	162.98	86.43
	15.49**	11.06**	19.33**	51.88**	36.14*
log per worker gdp (origin)	-22.83	6.12	-17.78	-75.07	-8.62
	20.25	10.21	13.13	23.52**	23.03
log distance	-31.23	-36.6	-29.45	-38.71	-142.64
	7.05**	7.60**	8.19**	12.23**	4.41**
common language	7.35	19.69	15.15	7.06	120.2
	8.39	12.93	11.45	12.61	9.18**
share of young population (origin)	65.14	217.99	-13.57	137.17	683.44
	99.75	114.24+	113.13	265.43	299.65*
log yrs schooling (destination)		-47.23	-32.23		
		15.04**	17.75+		
log yrs schooling (origin)		7.71	47.93		
		12.39	36.51		
log capital per worker (destination)			-27.91		
			12.58*		
log capital per worker (origin)			10.83		
			13.63		
constant	-0.61	-90.28	48.56	-862.22	62.86
	242.82	126.48	93.73	479.60+	406.78
number of obs	7411	7313	4103	1933	8010
R-squared	0.23	0.25	0.26	0.69	

Destination and origin countries' dummy variables are included in each specification. Standard errors, clustered by country pairs, are presented under each estimated coefficient. + significant at 10%; * significant at 5%; ** significant at 1%

The *emigration rate* (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) gives the number of incoming immigrants per 100,000 individuals in the origin country's population. In regression (1), I use terms of trade (lagged by one year) as an instrument for per worker GDP (lagged by one year) in both destination and origin country. The F value for the excluded instruments (terms of trade) is equal to 30.47 and 70.36 in the first stage of per worker GDP of destination and origin countries, respectively.

per worker gdp is the level of per worker GDP, PPP-adjusted (constant 1996 international dollars), lagged by one year.

distance is the great-circle distance. *common language* is a dummy variable equal to one if a common language is spoken in both destination and origin countries. *share of young population (origin)* is the share of the population in the origin country aged 15-29. *log yrs schooling* is the log of the average schooling years in the total population over age 15, lagged by one year. *log capital per worker* is non-residential capital stock per worker (1985 intl. prices), lagged by one year. See Appendix 2 for data sources.

Appendix 1. Average yearly inflows into each destination country, by country of origin (1980-1995)

Australia (1983-1995)

country of origin	inflow
UK	17095
New Zealand	11045
Vietnam	8048
Hong Kong	5739
Philippines	5379
Malaysia	3493
India	3069
China	2934
Former Yugoslavia	2790
South Africa	2441
Sri Lanka	2146
Lebanon	2089
USA	1724
Fiji	1682
Poland	1608
Ireland	1462
Chinese Taipei	1358
Germany	1303
Former USSR	1021
Portugal	767

percentage change (1983-1995) -6.22%

Belgium (1984-1995)

country of origin	inflow
France	6072
Netherlands	6014
USA	2930
Germany	2916
UK	2899
Morocco	2801
Italy	2495
Turkey	2239
Zaire	1966
Portugal	1435
Japan	833
Spain	833
Former Yugoslavia	829
Greece	759
Poland	655
China	589
Algeria	382
Tunisia	310

percentage change (1980-1995) 13.46%

Canada (1980-1995)

country of origin	inflow
Hong Kong	19334
India	10437
Philippines	9441
UK	9034
Vietnam	8791
Poland	7550
USA	7459
China	6292
Lebanon	3917
Sri Lanka	3791
Portugal	3653
Jamaica	3543
Chinese Taipei	3255
Guyana	3108
El Salvador	2697
Haiti	2243
Iran	2193
France	2070
Former Yugoslavia	1933
South Korea	1584
Trinidad Tobago	1433
Romania	1241
Pakistan	1037
Former USSR	791
Somalia	195

percentage change (1980-1995) 48.29%

Denmark (1990-1994)

country of origin	inflow
Somalia	1264
UK	1068
Turkey	1042
Germany	805
Iraq	789
Norway	699
Sweden	612
USA	606
Iran	570
Vietnam	549
Former Yugoslavia	481
Iceland	479
Poland	448
Thailand	366
Pakistan	356
Lebanon	335
Netherlands	304
France	269
Morocco	215
Italy	200
Finland	181

percentage change (1984-1994) 75.28%

The average yearly inflow into each destination country, by origin country, is calculated using only non-zero flow observations. Therefore, for some origin countries, the average of immigration inflows into the destination country is not taken over the years indicated at the top of each table, but over a shorter period of time.

percentage change is the percentage change of *total immigrant inflows* between 1980 and 1995 (for some destination countries, the percentage change is over a shorter period of time, as indicated). To calculate these percentage changes, I use directly aggregate data on total immigrant inflows. That is why the coverage of years differs from the summary statistics by country of origin.

Appendix 1. Average yearly inflows into each destination country, by country of origin (1980-1995) (cont.)

France (1984-1995)

country of origin	inflow
Morocco	11892
Algeria	9187
Turkey	5777
Tunisia	3083
Lebanon	2818
USA	2403
Haiti	2183
Portugal	2050
Vietnam	1761
Zaire	1437
Poland	1422
Japan	1219
China	1084
Former Yugoslavia	1084
Sri Lanka	899
Romania	891
Cambodia	860
Spain	400

percentage change (1980-1995) -6.23%

Germany* (1984-1995)

country of origin	inflow
Poland	117019
Former Yugoslavia	92124
Bosnia-Herzegovina	76836
Turkey	68791
Romania	61910
Italy	39184
Croatia	24056
Former USSR	23365
Hungary	21835
Greece	20372
Bulgaria	19245
USA	17670
Former CSFR	10692
Portugal	9654
Spain	4705
Morocco	4375
Slovenia	2658
Tunisia	2249

percentage change (1980-1995) 24.85%

Japan (1980-1995)

country of origin	inflow
China	35425
USA	35367
Philippines	35121
South Korea	21052
Chinese Taipei	10882
UK	9614
Brazil	6779
Hong Kong	6296
Thailand	5913
Germany	5334
Canada	3449
Peru	1008

percentage change (1980-1995) -42.10%

Luxembourg (1983-1995)

country of origin	inflow
Portugal	2170
France	1272
Belgium	897
Germany	662
Italy	441
Netherlands	281
USA	256
Spain	124

percentage change (1980-1995) 29.73%

The average yearly inflow into each destination country, by origin country, is calculated using only non-zero flow observations. Therefore, for some origin countries, the average of immigration inflows into the destination country is not taken over the years indicated at the top of each table, but over a shorter period of time.

* Figures for migrants from the former Yugoslavia to Germany do not include Croatia from 1992 and Bosnia-Herzegovina from 1993. Data from the former USSR to Germany does not include Russia from 1992.

percentage change is the percentage change of *total immigrant inflows* between 1980 and 1995 (for some destination countries, the percentage change is over a shorter period of time, as indicated). To calculate these percentage changes, I use directly aggregate data on total immigrant inflows. That is why the coverage of years differs from the summary statistics by country of origin.

Appendix 1. Average yearly inflows into each destination country, by country of origin (1980-1995) (cont.)

Netherlands (1984-1995)

country of origin	inflow
Turkey	8363
Former Yugoslavia	7392
Morocco	6537
Germany	5295
UK	4575
Suriname	4416
USA	2303
Belgium	2050
France	1517
Poland	1148
Italy	893

percentage change (1980-1995) -16.04%

Norway* (1984-1995)

country of origin	inflow
Bosnia-Herzegovina	3728
Denmark	2201
Sweden	1526
UK	1253
USA	987
Former Yugoslavia	868
Pakistan	682
Iran	669
Vietnam	612
Chile	537
Somalia	468
Sri Lanka	450
Germany	399

percentage change (1980-1995) 39.83%

Sweden (1984-1995)

country of origin	inflow
Bosnia-Herzegovina	16972
Iran	4048
Finland	3880
Norway	3118
Former Yugoslavia	2840
Iraq	2051
Denmark	1877
Somalia	1724
Chile	1631
Poland	1484
Turkey	1214
Ethiopia	947
Russian Federation	910
Lebanon	896
USA	831
Croatia	784
Germany	761
Romania	746
UK	715
Thailand	603
India	369
Greece	311

percentage change (1983-1995) 61.88%

Switzerland (1984-1995)

country of origin	inflow
Former Yugoslavia	18716
Portugal	9085
Germany	8333
Italy	8216
France	4655
Spain	4402
Turkey	4195
USA	2530
UK	2407
Austria	1728
Netherlands	1607
Canada	687

percentage change (1980-1995) 24.68%

The average yearly inflow into each destination country, by origin country, is calculated using only non-zero flow observations. Therefore, for some origin countries, the average of immigration inflows into the destination country is not taken over the years indicated at the top of each table, but over a shorter period of time.

* Figures for migrants from the former Yugoslavia to Norway do not include Bosnia-Herzegovina from 1993.

percentage change is the percentage change of *total immigrant inflows* between 1980 and 1995 (for some destination countries, the percentage change is over a shorter period of time, as indicated). To calculate these percentage changes, I use directly aggregate data on total immigrant inflows. That is why the coverage of years differs from the summary statistics by country of origin.

Appendix 1. Average yearly inflows into each destination country, by country of origin (1980-1995) (cont.)

United Kingdom (1982-1995)

country of origin	inflow
Pakistan	5817
India	5047
Bangladesh	3796
USA	3776
Australia	2659
New Zealand	1964
Nigeria	1556
Iran	1501
Japan	1474
Hong Kong	1287
Ghana	1093
Canada	1035
Sri Lanka	1021
Philippines	986
South Africa	926
Turkey	822
Jamaica	775
Malaysia	701
Iraq	500
Kenya	481
Poland	481
Thailand	444
Germany	419
Cyprus	402
Morocco	380
Spain	363
Sweden	355
France	345
Italy	340
Netherlands	289
Former Yugoslavia	276
Portugal	223

percentage change (1980-1995) -20.49%

United States (1980-1995)

country of origin	inflow
Mexico	199862
Philippines	51886
Vietnam	45041
China	32824
Dominican Republic	30471
India	29754
South Korea	29197
Former USSR	23231
El Salvador	21901
Jamaica	20219
Cuba	15174
Haiti	15168
UK	14939
Iran	14596
Poland	13534
Canada	12980
Chinese Taipei	12962
Colombia	12696
Laos	12165
Ireland	12054
Guatemala	9328
Guyana	9243
Cambodia	8108
Pakistan	7725
Peru	7637
Germany	7005
Hong Kong	6994
Thailand	6270
Ecuador	6189
Nicaragua	5626
Honduras	5507
Bangladesh	2684

percentage change (1980-1995) 35.79%

The average yearly inflow into each destination country, by origin country, is calculated using only non-zero flow observations. Therefore, for some origin countries, the average of immigration inflows into the destination country is not taken over the years indicated at the top of each table, but over a shorter period of time.

percentage change is the percentage change of *total immigrant inflows* between 1980 and 1995 (for some destination countries, the percentage change is over a shorter period of time, as indicated). To calculate these percentage changes, I use directly aggregate data on total immigrant inflows. That is why the coverage of years differs from the summary statistics by country of origin.

Appendix 2. Summary Statistics (1980-1995) and Data Sources

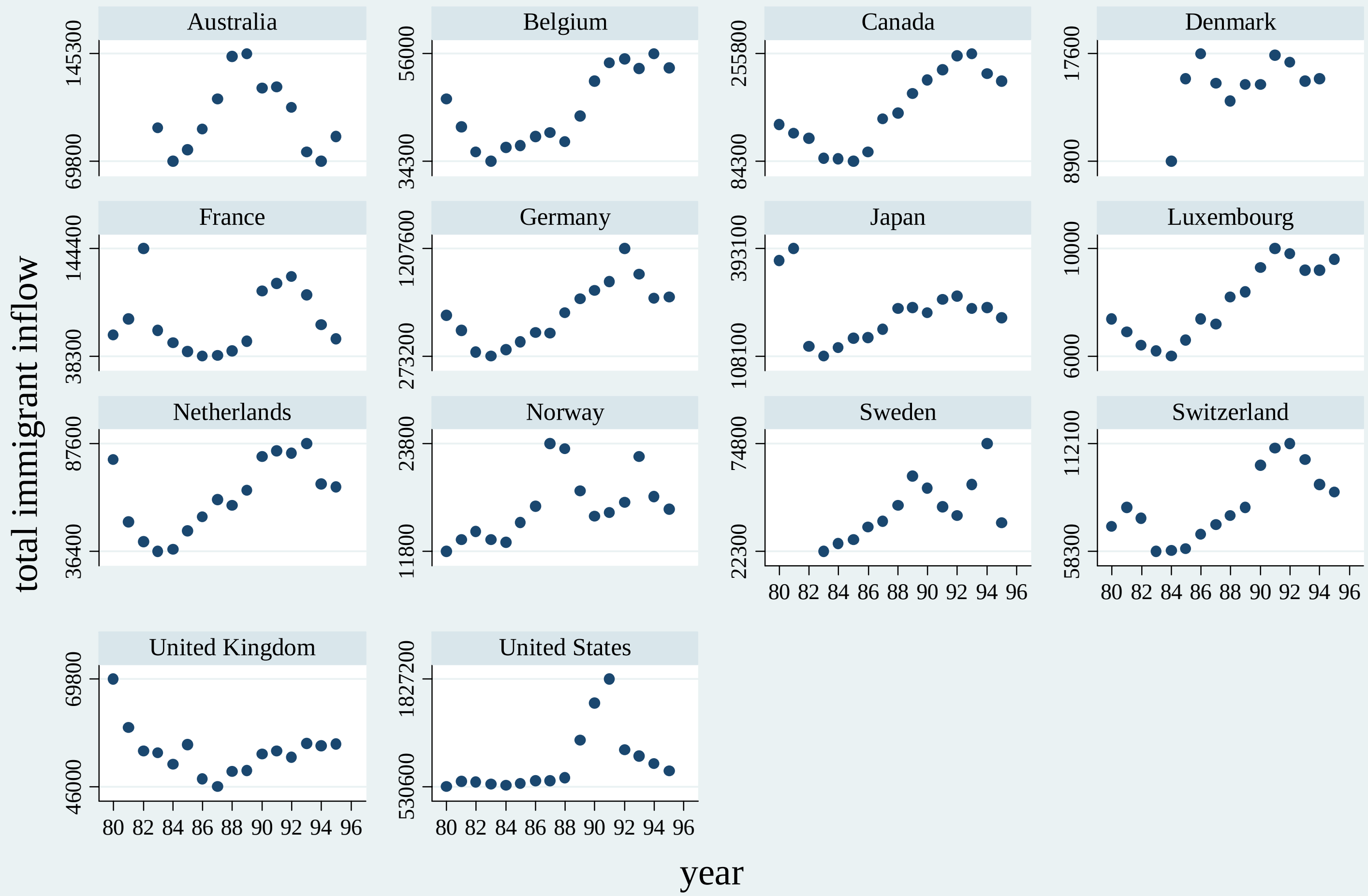
Variable	Obs	Mean	Std. Dev.	Min	Max
emigration rate	8010	13.2433	81.5410	0.0000	1568.9430
per worker gdp (destination)	8010	40682	5895	25252	55361
per worker gdp (origin)	8010	20061	14106	1027	55361
log distance	8010	8.1715	0.8694	5.0872	9.3836
land border	8010	0.0268	0.1616	0	1
common language	8010	0.1704	0.3760	0	1
colony	8010	0.0385	0.1923	0	1
share of young population (origin)	8010	0.2612	0.0303	0.1951	0.3152
years schooling (destination)	4103	9.6403	1.3096	6.8370	11.8650
years schooling (origin)	4103	7.0285	2.4659	2.7240	11.8650
capital per worker (destination)	4103	36041	12167	16992	76733
capital per worker (origin)	4103	19232	13290	822	48135
unemployment rate (destination)	5010	6.7306	3.4646	0.5000	14.1000
unemployment rate (origin)	5010	8.1476	5.2840	0.0800	27.6000
origin country's relative inequality	4028	1.2123	0.3846	0.3861	2.6810

The *emigration rate* (immigrant inflow from origin to destination country (multiplied by 100,000), divided by origin country's population) is from the IMS data set (OECD 1997). *Per worker GDP*, PPP-adjusted (constant 1996 international dollars) is from the Penn World Tables, version 6.1. *Log distance*, *land border*, *common language*, and *colony* (countries ever in a colonial relationship) are from Glick and Rose (2001). *Years of schooling* are from Barro and Lee (2000) data set. *Capital per worker* (Nonresidential Capital Stock per Worker, 1985 intl. prices) is from the Penn World Tables, version 5.6.

The share of young population (origin) is based on data from the United Nations. The unemployment rate is from the World Development Indicators (2001), World Bank. The *origin country's relative inequality* is based on data on Gini coefficients from Deininger and Squire (1996) data set (only high-quality observations were used). Data on terms of trade comes from the World Bank's Global Development Network Growth Database, Macro Time Series (2002). The data set on immigration policy changes was constructed by Anna Maria Mayda and Krishna Patel (2004) - see Appendix to the paper. All time-varying variables (except the emigration rate) are lagged by one year.

Summary statistics for the *emigration rate*, *per worker gdp (destination), per worker gdp (origin), log distance*, *land border*, *common language*, *colony, share of young population (origin)* are based on the same observations as in the pooled regressions in Table 1. Summary statistics for *years schooling (destination)*, *years of schooling (origin)*, *capital per worker (destination)*, *capital per worker (origin)* are based on the same observations as in regression (3), Table 3. Summary statistics for *unemployment rate (destination)* and *unemployment rate (origin)* are based on the same observations as in regression (4), Table 2. Finally, summary statistics for the *origin country's relative inequality* are based on the same observations as in regression (1), Table 2.

Figure 1: Total immigrant inflow by destination country



Graphs by country of destination

Figure 2: The actual emigration rate as a function of mean income opportunities in the destination and origin country and of mean moving costs